SELECTION IN FACTORIAL EXPERIMENTS

Robert E. Bechhofer
School of Operations Research and Industrial Engineering
Cornell University

ABSTRACT

The performance of many if not most real-life systems which are simulated depends on two or more factors which can be set at various "levels." In order to understand the behavior of such a system the experimenter must conduct a factorial experiment in which the behavior of the system is studied at selected factor-level combinations. It is often of interest to select the "best" factor-level combination, i.e., the one associated with the highest average response.

The purpose of this paper is to discuss some of the more commonly used statistical selection procedures. Selection procedures appropriate for use with single-factor experiments are considered first. The ideas associated with the use of these procedures are then generalized to two-factor experiments (and implicitly to multi-factor experiments) and new selection procedures for use with such multi-factor experiments are described.

1. INTRODUCTION AND SUMMARY

In an interesting paper presented at the 1976 Winter Simulation Conference, Dudewicz [1976] described some applications of ranking and selection procedures in simulation studies. Earlier, Kleijnen [1975] (see, in particular, pp. 553-561 and 599-675) had given a broad discussion of the virtues and drawbacks of a large variety of such procedures with particular reference to their use in simulation studies. Dudewicz [1976] dealt exclusively with single-factor experiments, i.e., experiments in which the individual conducting the simulation varies only a single factor, e.g., job shop precedence rule, in order to determine which such rule yields (say) the highest average output. However, the performance of many if not most real-life systems which are simulated depends not on one factor but on two or more factors which can be set at various "levels." For example, the experimenter may choose to vary not only the job shop precedence rule but also (and simultaneously) the form of the arrival time distribution in order to study their joint effect on the average output of the system. This type of problem is mentioned by Kleijnen (p. 561). Thus, in order to understand the behavior of such a system, the experimenter must actually conduct a factorial experiment in which the behavior of the system is studied at selected factor-level combinations. It is with such multi-factor experiments that we will be concerned in the present paper.

It is our purpose in this paper to provide the reader with relevant background concerning some of the more commonly used selection procedures. First we discuss single-factor experiments--in particular, we state precisely the statistical assumptions that are usually made, describe two of the now classical statistical formulations of the selection problem, and give literature references to several of the procedures which provide solutions to these problems. We next discuss two-factor experiments (and implicitly multi-factor experiments), and show how the single-factor statistical assumptions and formulations used with selection problems can be generalized in a natural way to deal with corresponding multi-factor selection problems. We then present newly-devised selection procedures which provide solutions for these multi-factor selection problems.

The new procedures follow from as yet unpublished research undertaken jointly by the author and Professor Charles W. Dunnett, Department of Clinical Epidemiology and Biostatistics, McMaster University, Ontario, Canada. The theory underlying the procedures reported herein, and related procedures, is developed in papers presently being prepared.

2. SINGLE-FACTOR EXPERIMENTS

2.1 STATISTICAL ASSUMPTIONS

We assume that we have k populations Π_i $(1 \leq i \leq k)$ of normally distributed data, the ith population having population mean μ_i and population variance σ_i^2; population Π_i $(1 \leq i \leq k)$ should be thought of as being associated with the ith "level" of a qualitative factor. The μ_i are assumed to be unknown. Let $\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}$ denote the ranked values of the μ_i; it is assumed that the pairing of the Π_i with the $\mu_{[j]}$ $(1 \leq i,j \leq k)$ is completely unknown. In this exposition we also assume that $\sigma_i^2 = \sigma^2$ $(1 \leq i \leq k)$, the common value being assumed known or being assumed unknown. The general case in which the values of the σ_i^2 are assumed known but perhaps unequal, or are assumed completely unknown will not be considered here; for the former see Bechhofer [1954], p. 24, and for the latter see Dudewicz and Dalal [1975] or Rinott [1974]. We denote the mth observation from Π_i by X_{im} $(1 \leq i \leq k;\ m = 1,2,\ldots)$, all observations being assumed independent.

2.2 TWO FORMULATIONS OF THE SELECTION PROBLEM

The two most commonly used formulations of the selection problem are due to Bechhofer [1954] and Gupta [1956], [1965]; these are referred to as the indifference-zone approach and the subset approach, respectively. The formulations are described below.

2.2.1 The Indifference-Zone Approach

The goal and probability requirement associated with the indifference-zone approach are:

Goal: "To select the level (population) associated with $\mu_{[k]}$." (2.1)

It is assumed that prior to the start of experimentation the experimenter can specify two constants $\{\delta^*, P^*\}$ $(0 < \delta^* < \infty,\ 1/k < P^* < 1)$ which are then incorporated into the following probability requirement:

Probability requirement:

Prob{Selecting the level (population) associated with $\mu_{[k]}\} \geq P^*$ whenever $\mu_{[k]} - \mu_{[k-1]} \geq \delta^*$. (2.2)

The experimenter then restricts consideration to procedures which guarantee (2.2).

2.2.2 The Subset Approach

The goal and probability requirement associated with the subset approach are:

Goal: "To select a (non-empty) subset of the levels (populations) which contains the level (population) associated with $\mu_{[k]}$." (2.3)

It is assumed that prior to the start of experimentation the experimenter can specify a constant $\{P^*\}$ $(1/k < P^* < 1)$ which is then incorporated into the following probability requirement:

Probability requirement:

Prob{Selecting a subset of the levels (populations) which contains the level (population) associated with $\mu_{[k]}\} \geq P^*$ regardless of the values of the μ_i $(1 \leq i \leq k)$. (2.4)

The experimenter then restricts consideration to procedures which guarantees (2.4).

The goal (2.3) can be thought of as a screening goal which is particularly appropriate when the number of levels (populations) is large. Having conducted an experiment employing (2.3), the experimenter might follow up with an experiment employing (2.1). See Remark 2.1.

2.3 PROCEDURES

2.3.1 The Indifference-Zone Approach

If the experimenter wishes to guarantee (2.2), and the value of the common variance σ^2 is assumed known, then he can use the single-stage procedure of Bechhofer [1954]; this procedure is described in Dudewicz [1976], Section II, and Kleijnen [1975], pp. 601-607.

If the experimenter wishes to guarantee (2.2), and the value of the common variance is assumed unknown, then he can use the two-stage procedure of Bechhofer, Dunnett and Sobel [1954] (but he cannot use a single-stage procedure--see Dudewicz [1971]); this procedure is described in Kleijnen [1975], pp. 608-610.

2.3.2 The Subset Approach

If the experimenter wishes to guarantee (2.4), and the value of the common variance is assumed known or is assumed unknown, then he can use the single-stage procedure of Gupta [1956], [1965]; this procedure is described in Kleijnen [1975], p. 555.

Remark 2.1: Recently Tamhane and Bechhofer [1977] proposed a two-stage procedure which guarantees (2.2) when the value of the common variance is assumed known. This two-stage procedure has the highly desirable property that the expected total number of observations required by the procedure is always less than the total number of observations required by the corresponding single-stage procedure of Bechhofer [1954], regardless of the configuration of the population means. The two-stage procedure can be regarded as a composite one which performs a screening function in the first stage, and selects a best level (population) in the second stage from among those levels (populations) not screened out in the first stage.

3. TWO-FACTOR EXPERIMENTS

3.1 STATISTICAL ASSUMPTIONS

We assume that we have $r \times c$ populations Π_{ij} $(1 \leq i \leq r,\ 1 \leq j \leq c)$ of normally distributed data, the (i,j)th population having population mean μ_{ij} and population variance σ_{ij}^2; population Π_{ij} $(1 \leq i \leq r,\ 1 \leq j \leq c)$ should be thought of as being associated with the ith "level" of the first qualitative factor and the jth "level" of the second qualitative factor. The μ_{ij} are assumed to be unknown. We write

$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$ $(1 \leq i \leq r,\ 1 \leq j \leq c)$

where $\sum_{i=1}^{r} \alpha_i = \sum_{j=1}^{c} \beta_j = \sum_{i=1}^{r} \gamma_{ij} = \sum_{j=1}^{c} \gamma_{ij} = 0.$

Then α_i is referred to as the "effect" on the population mean of the ith level of Factor A, β_j is referred to as the "effect" on the population mean of the jth level of Factor B, and γ_{ij} is referred to as the joint "effect" on the population mean of the ith level of Factor A and the jth level of Factor B. The quantity γ_{ij} is also referred to as the first-order (or two-factor) interaction "effect."

If $\gamma_{ij} \not\equiv 0$ (all i,j) then we say that interaction exists between the levels of the two factors. In this situation it usually is not meaningful to seek the "best" level of the first factor and the "best" level of the second factor (since each depends on the level of the other factor); however, it is meaningful to seek the "best" factor-level combination. To this end we let $\mu_{[1]} \leq \mu_{[1]} \leq \cdots \leq \mu_{[rc]}$ denote the ranked values of the μ_{ij}. It is assumed that the pairing of the Π_{ij} with the $\mu_{[p]}$ $(1 \leq i \leq r,\ 1 \leq j \leq c;\ 1 \leq p \leq rc)$ is completely unknown.

If $\gamma_{ij} \equiv 0$ (all i,j) then we say that no interaction exists between the levels of the two factors. Here it is meaningful to seek the "best" level of the first factor and the "best" level of the second factor. We thus let $\alpha_{[1]} \leq \alpha_{[2]} \leq \cdots \leq \alpha_{[r]}$ and $\beta_{[1]} \leq \beta_{[2]} \leq \cdots \leq \beta_{[c]}$ denote the ranked values of the α_i and β_j, respectively. It is assumed that the pairing of the Π_{ij} with the $\alpha_{[p]}$ and/or the $\beta_{[q]}$ $(1 \leq i \leq r,\ 1 \leq j \leq c;\ 1 \leq p \leq r,\ 1 \leq q \leq c)$ is completely unknown.

We also assume that $\sigma_{ij}^2 = \sigma^2$ $(1 \leq i \leq r,\ 1 \leq j \leq c)$, the common value being assumed known or being assumed unknown. We denote the mth observation from Π_{ij} by X_{ijm} $(1 \leq i \leq r;\ 1 \leq j \leq c;\ m = 1,2,\ldots)$, all observations being assumed independent.

3.2 TWO FORMULATIONS OF THE SELECTION PROBLEM

As we did in Section 2.2 for single-factor experiments, we now describe the use of the indifference-zone approach and the subset approach for two-factor experiments.

3.2.1 The Indifference-Zone Approach

3.2.1.1 Interaction Between the Levels of the Factors. The goal and probability requirement associated with the indifference-zone approach for $\gamma_{ij} \not\equiv 0$ (all i,j) are the same as (2.1) and (2.2) with k replaced by rc.

3.2.1.2 No Interaction Between the Levels of the Factors. For $\gamma_{ij} \equiv 0$ (all i,j), the goal and probability requirement are:

Goal: "To select the level of Factor A associated with $\alpha_{[r]}$, and simultaneously to select the level of Factor B associated with $\beta_{[c]}$." (3.1)

It is assumed that prior to the start of experimentation the experimenter can specify three constants $\{\delta_\alpha^*, \delta_\beta^*, P^*\}$ $(0 < \delta_\alpha^*, \delta_\beta^* < \infty,\ 1/rc < P^* < 1)$ which are then incorporated into the following probability requirement:

Probability requirement:

Prob{Selecting the level of Factor A associated with $\alpha_{[r]}$, and simultaneously of selecting the level of Factor B associated with $\beta_{[c]}\} \geq P^*$

$$\text{whenever } \begin{cases} \alpha_{[r]} - \alpha_{[r-1]} \geq \delta_\alpha^* \\ \text{and} \\ \beta_{[c]} - \beta_{[c-1]} \geq \delta_\beta^*. \end{cases} \quad (3.2)$$

3.2.2 The Subset Approach

3.2.2.1 Interaction Between the Levels of the Factors. The goal and probability requirement associated with the subset approach for $\gamma_{ij} \not\equiv 0$ (all i,j) are the same as (2.3) and (2.4) with k replaced by rc.

3.2.2.2 No Interaction Between the Levels of the Factors. For $\gamma_{ij} \equiv 0$ (all i,j), the goal and probability requirement are:

Goal: "To select a (non-empty) subset of the levels of Factor A which contains the level associated with $\alpha_{[r]}$, and simultaneously, to select a (non-empty) subset of the levels of Factor B which contains the level associated with $\beta_{[c]}$." (3.3)

It is assumed that prior to the start of experimentation the experimenter can specify a constant $\{P^*\}$ $(1/rc < P^* < 1)$ which is then incorporated into the following probability requirement:

Probability requirement:

Prob{Selecting a subset of the levels of Factor A which contains the level associated with $\alpha_{[r]}$, and simultaneously, of selecting a subset of the levels of Factor B which contains the level associated with $\beta_{[c]}\} \geq P^*$ (3.4)

regardless of the values of the α_i and β_j $(1 \leq i \leq r,\ 1 \leq j \leq c)$.

3.3 PROCEDURES

In this section we assume $\gamma_{ij} \equiv 0$ (all i,j), and give procedures which will guarantee (3.2) and (3.4) when the value of the common variance σ^2 is assumed known, and when the value of the common variance is assumed unknown. The procedure are generalizations of procedures previously proposed for single-factor experiments.

3.3.1 The Indifference-Zone Approach

If the experimenter wishes to guarantee (3.2), and the value of the common variance is assumed known, then he can use the single-stage procedure of Bechhofer [1954], Section 4 (and Example 3, p. 37); see also, Kleijnen [1975], pp. 634-636.

The virtue of conducting one two-factor experiment rather than two independent single-factor experiments to guarantee (3.2) is discussed by Bawa [1972]. In effect the factorial design of the experiment makes the data "work twice" and is in this sense more efficient than two independent single-factor experiments; this results in a saving (sometimes substantial) in the total number of observations required to guarantee (3.2).

If the experimenter wishes to guarantee (3.2), and the value of the common variance is assumed unknown, then he can use the following new two-stage procedure of Bechhofer and Dunnett [1977] (which is a generalization of the two-stage procedure of Bechhofer, Dunnett and Sobel [1954]); constants g (which depend on $\{(r,c),\nu,\delta^*_\alpha/\delta^*_\beta,P^*\}$) are given in B&D [1977]):

"a) In the first stage take an arbitrary common number $N_0 > 1$ of observations from each of the rc populations Π_{ij} $(1 \leq i \leq r, 1 \leq j \leq c)$.

b) Calculate

$$S^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \sum_{m=1}^{N_0} (X_{ijm} - \sum_{m=1}^{N_0} X_{ijm}/N_0)^2/\nu$$ which

is an unbiased estimate of σ^2 based on $\nu = rc(N_0-1)$ degrees of freedom.

c) Enter the appropriate table (e.g., abbreviated Table 1, below, for $(r,c) = (2,2)$ and $(2,3)$, selected ν, $\delta^*_\alpha/\delta^*_\beta = 1$, $P^* = 0.95$) and obtain a constant g.

d) In the second stage, take a common number $N-N_0$ of additional observations from each of the k populations where

$$N = N_0 \quad \text{if} \quad M < N_0$$
$$N = [M] \quad \text{if} \quad M > N_0 \qquad (3.5)$$

where $M = (\sqrt{2}gS/\delta^*_\beta)^2/r$ and $[M]$ denotes the smallest integer equal to or greater than M.

e) Calculate the $r+c$ over-all (first stage plus second-stage) sample sums $\sum_{j=1}^{c} \sum_{m=1}^{N} X_{ijm}$ $(1 \leq i \leq r)$, $\sum_{i=1}^{r} \sum_{m=1}^{N} X_{ijm}$ $(1 \leq j \leq c)$, and let

$$\sum_{j=1}^{c} \sum_{m=1}^{N} X_{[r]jm} = \max\{\sum_{j=1}^{c} \sum_{m=1}^{N} X_{ijm} \ (1 \leq i \leq r)\},$$
$$\sum_{i=1}^{r} \sum_{m=1}^{N} X_{i[c]m} = \max\{\sum_{i=1}^{r} \sum_{m=1}^{N} X_{ijm} \ (1 \leq j \leq c)\}.$$

f) Select the level that yielded $\sum_{j=1}^{c} \sum_{m=1}^{N} X_{[r]jm}$ as the one associated with $\alpha_{[r]}$, and the level that yielded $\sum_{i=1}^{r} \sum_{m=1}^{N} X_{i[c]m}$ as the one associated with $\beta_{[c]}$."

TABLE 1

Values of g for $P^* = 0.95$, $\delta^*_\alpha/\delta^*_\beta = 1$ and $(r,c) = (2,2)$ and $(2,3)$.

ν	r = 2, c = 2	ν	r = 2, c = 3
4	2.7215	6	2.4939
8	2.2849	12	2.2155
12	2.1645	18	2.1340
16	2.1083	24	2.0952
20	2.0759	30	2.0713
24	2.0548	36	2.0565
28	2.0392	42	2.0460
32	2.0276		
36	2.0192		
40	2.0125		
∞	1.9545	∞	1.9857

The values in this table are abstracted from tables in Bechhofer and Dunnett [1977] which give many additional g-values for selected $\{(r,c),\nu,\delta^*_\alpha/\delta^*_\beta,P^*\}$.

Remark 3.1: The g-entry associated with $\nu = \infty$ in Table 1 can also be used if the experimenter wishes to guarantee (3.2), and the value of the common variance is assumed known (which is implied if $\nu = \infty$). Then a single-stage procedure is used with a common number $N = [(\sqrt{2}g\sigma/\delta^*_\beta)^2/r]$ of observations being taken from each of the rc populations Π_{ij} $(1 \leq i \leq r, 1 \leq j \leq c)$; the decision rule is the same as that given in step f) of (3.5) with the phrase in e) "first stage plus second stage" replaced by "single-stage." (This rule is then equivalent to the one given in Bechhofer [1954], Section 4.)

3.3.2 The Subset Approach

If the experimenter wishes to guarantee (3.4), and the value of the common variance is assumed known or is assumed unknown, then he can use the following new single-stage procedure of Bechhofer and Dunnett [1977] (which generalizes the single-stage procedure of Gupta [1956], [1965]); constants h depending on $\{(r,c),\nu,P^*\}$ are given in B&D [1977]).

"a) Take an arbitrary common number $N > 1$ of observations from each of the rc populations Π_{ij} $(1 \leq i \leq r,\ 1 \leq j \leq c)$.

b) Calculate

$$S^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{m=1}^{N}(X_{ijm} - \sum_{m=1}^{N} X_{ijm}/N)^2/\nu \quad \text{which}$$

is an unbiased estimate of σ^2 based on $\nu = rc(N-1)$ degrees of freedom.

c) Enter the appropriate table (e.g., abbreviated Table 2, below, for $(r,c) = (2,2)$ and $(2,3)$, selected ν, $P^* = 0.95$) and obtain a constant h. (3.6)

d) Calculate the $r+c$ sample means

$$\overline{X}_{i\cdot\cdot} = \sum_{j=1}^{c}\sum_{m=1}^{N} X_{ijm}/cN \quad (1 \leq i \leq r),$$

$$\overline{X}_{\cdot j\cdot} = \sum_{i=1}^{r}\sum_{m=1}^{N} X_{ijm}/rN \quad (1 \leq j \leq c), \quad \text{and let}$$

$$\overline{X}_{[r]\cdot\cdot} = \max\{\overline{X}_{i\cdot\cdot}\ (1 \leq i \leq r)\},$$

$$\overline{X}_{\cdot[c]\cdot} = \max\{\overline{X}_{\cdot j\cdot}\ (1 \leq j \leq c)\},$$

e) Retain the ith level of Factor A $(1 \leq i \leq r)$ among the selected levels of Factor A if and only if

$$\overline{X}_{i\cdot\cdot} \geq \overline{X}_{[r]\cdot\cdot} - \sqrt{2}hS/\sqrt{cN},$$

and retain the jth level of Factor B $(1 \leq j \leq c)$ among the selected levels of Factor B if and only if

$$\overline{X}_{\cdot j\cdot} \geq \overline{X}_{\cdot[c]\cdot} - \sqrt{2}hS/\sqrt{rN}."$$

TABLE 2

Values of h for $P^* = 0.95$ and $(r,c) = (2,2)$ and $(2,3)$.

ν	$r = 2$, $c = 2$	ν	$r = 2$, $c = 3$
4	2.7215	6	2.6469
8	2.2849	12	2.3479
12	2.1645	18	2.2603
16	2.1083	24	2.2186
20	2.0759	30	2.1930
24	2.0548	36	2.1770
28	2.0392	42	2.1658
32	2.0276		
36	2.0192		
40	2.0125		
∞	1.9545	∞	2.1009

The values in this table are abstracted from tables in Bechhofer and Dunnett [1977] which give many additional h-values for selected $\{(r,c),\nu,P^*\}$.

Note: If $r = c$ then g (in Table 1) equals h (in Table 2).

4. MULTI-FACTOR EXPERIMENTS

No new ideas are encountered for multi-factor selection problems (i.e., selection problems arising from experiments involving three or more qualitative factors) which were not already present in two-factor selection problems. The method of generalization from two to three or more factors is clear; additional tables of the g- and h-constants are necessary; some of these will be contained in Bechhofer and Dunnett [1977].

5. CONCLUDING REMARKS

The reader who is interested in further study of selection procedures is referred to Gibbons, Olkin, and Sobel [1977], and a textbook in preparation by S. S. Gupta and S. Panchapakesan.

6. ACKNOWLEDGMENT

Research supported in part by contracts DAAG29-77-C-0003, U.S. Army Research Office-Durham; and N00014-75-C-0586, Office of Naval Research. Approved for public release; distribution unlimited.

7. BIBLIOGRAPHY

Bawa, V.S. (1972): "Asymptotic efficiency of one R-factor experiment relative to R one-factor experiments for selecting the best normal population," Journal of the American Statistical Association, Vol. 67, No. 339, pp. 660-661.

Bechhofer, R.E. (1954): "A single-sample multiple-decision procedure for ranking means of normal populations with known variances," Annals of Mathematical Statistics, Vol. 25, No. 1, pp. 16-39.

Bechhofer, R.E. and Dunnett, C.W. (1977): "Selecting the best factor-level combination associated with normal means arising from a factorial experiment when the common variance is unknown." Paper in preparation.

Bechhofer, R.E., Dunnett, C.W. and Sobel, M. (1954): "A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance," Biometrika, Vol. 41, Parts 1 and 2, pp. 170-176.

Dudewicz, E.J. (1971): "Nonexistence of a single-sample selection procedure whose P(CS) is independent of the variances," South African Statistical Journal, Vol. 5, pp. 37-39.

Dudewicz, E.J. (1976): "Statistics in simulation: how to design for selecting the best alternative," Proceedings of the 1976 Bicentenial Winter Simulation Conference (ed. by H. J. Highland, T. J. Schriber, and R. G. Sargent), pp. 67-71.

Dudewicz, E.J. and Dalal, S.R. (1975): "Allocation of observations in ranking and selection with unequal variances," Sankhya, B, Vol. 37, Part 1, pp. 28-78.

Gibbons, J.D., Olkin, I., and Sobel, M. (1977): Selecting and Ordering Populations (a new statistical methodology). Wiley-Interscience, New York.

Gupta, S.S. (1956): "On a decision rule for a problem in ranking means," Institute of Statistics Mimeograph Series No. 150, University of North Carolina, Chapel Hill, N.C.

Gupta, S.S. (1965): "On some multiple decision (selection and ranking) rules," Technometrics, Vol. 7, No. 25, pp. 225-245.

Kleijnen, J.P.C. (1975): Statistical Techniques in Simulation, Part II, Marcel Dekker, Inc., New York.

Rinott, Y. (1974): "On two stage procedures for selecting the population with the largest mean from several normal populations with unknown variances," Report, Department of Mathematics, Cornell University. Submitted for publication.

Tamhane, A.C. and Bechhofer, R.E. (1977): "A two-stage minimax procedure with screening for selecting the largest normal mean," Communications in Statistics, A, Vol. 6, Issue 11.